Accuracy And Efficiencies Associated With AI-Based Linking Of Clinical Trial Publications In Systematic Reviews

Md Sohail Aman, M.Pharm[1], Abhra R. Choudhury, MDS[1], Kopal Dixit, MSc[1], Anup Shaw, BTech[1], Iain Fotheringham, BSc[2];
[1]PharmaQuant Insights Pvt. Ltd., Kolkata, India, [2]PharmaQuant International Limited, Winchester, United Kingdom

PharmaQuant
Focused Insights, Better Outcomes

MSR82

BACKGROUND

- Systematic literature reviews (SLRs) frequently identify multiple publications originating from the same clinical study. Accurate linking of these related publications to their parent clinical trial is essential to ensure the integrity of evidence synthesis.[1] Failure to correctly link publications may result in duplicate data extraction, increased reviewer burden, inconsistencies in study-level datasets, and potential double-counting during meta-analysis or health technology assessment (HTA) submissions.[2]
- Manual clinical trial publication linking is time-consuming, and more so in large and complex evidence bases.[3] Reviewers must often rely on manual verification of trial identifiers, study characteristics, etc., which increases the risk of human error and variability in reviewer performance.[1]
- Recent artificial intelligence (AI) advances have enabled automated identification and linkage of related clinical trial publications. AI-supported tools could potentially improve scientific evidence traceability, standardize linking workflows, and reduce operational timelines in evidence generation.
- ActiveSLR® is a free-to-use, AI-enabled literature review platform designed to streamline literature screening, linking, and data extraction workflows.[4] In this analysis, ActiveSLR® was used as an example to attempt to quantify potential time savings associated with AI-supported study linking.

OBJECTIVES

The objective of this study was to evaluate the accuracy and operational efficiency of AI-supported publication linking using the ActiveSLR® linking feature as an example compared with conventional manual linking approaches performed in spreadsheet-based workflows.

METHODS

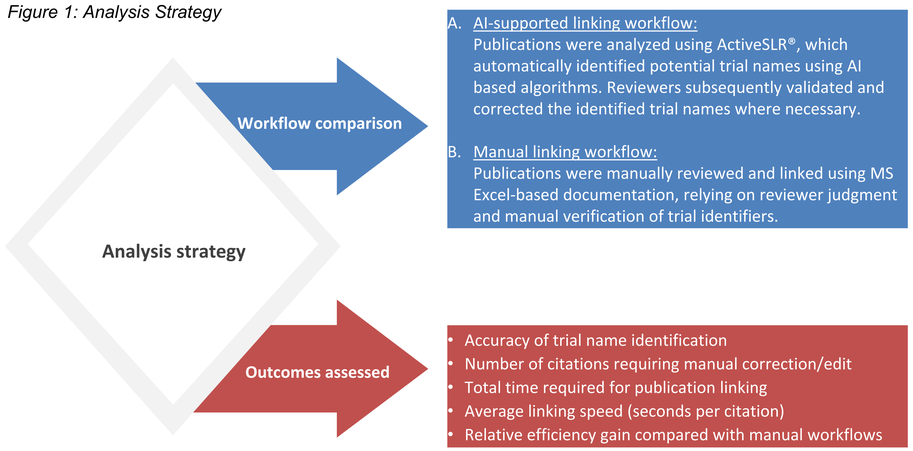
An evaluation was conducted using data from two independent SLRs with citations from MEDLINE and Embase in distinct therapeutic areas. Publication linking workflows were assessed using both AI-supported and manual approaches.

SLR Details:

Two independent SLR projects with citations from MEDLINE and Embase were included:

- **Project 1**: An SLR conducted in November 2025, on the clinical effectiveness of treatments in a neurodegenerative disease focusing on randomized controlled trials (RCTs) reporting overall survival (OS), progression-free survival (PFS), and objective response rate (ORR).
- **Project 2:** An SLR conducted in May 2025, on the clinical effectiveness and safety of treatments in a rare disease focusing on RCTs reporting OS, PFS, ORR, and any occurrence of grade 3-5 and/or selected adverse events.

Figure 1: Analysis Strategy



RESULTS

Study Volume and Citation Flow

A total of **4,197** citations were identified during database searches across both SLRs. Of these:

- A total of **179** citations were included in the final evidence base and assessed for publication linking. The findings of this analysis is presented in Table 1.
- Among included publications, ActiveSLR®'s AI-based linking correctly identified trial names for 164 citations, representing ~92% accuracy.
- Manual edits or corrections were made on 14 citations across both projects as these were abstract booklets, containing multiple in single PDFs.
- AI-based linking correctly identified the trial names of 164/165 full-text citations, representing ~99.4% accuracy.
- AI-supported reviewer validation and editing time both averaged under 10 seconds per citation.
- On average, linking time across both projects (including validation and edits) was ~24.6 seconds per citation.
- Manual linking in Excel required 3.82 minutes per citation on average, corresponding to an overall efficiency gain of ~89.3% in via AI-based linking.
- Subsequently linked publications helped identify those relevant for data extraction, leading to a 17.5% efficiency gain.

Table 1: Findings

Metric	ActiveSLR®	MS-Excel	ActiveSLR®	MS-Excel	Overall ActiveSLR®	Overall Excel
	Project 1		Project 2			
Citation volume screened	1,497	1,497	2,700	2,700	4,197	4,197
Citations evaluated for linking	53	53	126	126	179	179
Correctly identified trial names	41	53	123	125	164	178
Manual editing required	11	0	3	0	14	0
Incorrect trial name	1	0	0	1	1	1
Accuracy, %	77%	100%	98%	99%	92%	99%
Reviewer validation time	8.44 min	—	20.3 min	—	28.74 min	—
Average time per citation	24.2 sec	3 min	24.8 sec	4.16 min	24.6 sec	3.82 min
Total time required	21.37 min	159 min	52.08 min	524.78 min	73.45 min	683.78 min
Time taken to select studies eligible for extraction	36.2 min	39.75 min	86.1 min	108.42 min	122.31 min	148.2 min
Efficiency gain vs Excel for linking	**89.30%**					
Efficiency gain vs Excel for choosing eligible studies for data extraction	**17.5%**					

CONCLUSION

- ActiveSLR®'s AI-based linking demonstrated high accuracy in identifying and linking clinical trial publications across two SLRs, correctly identifying trial names for approximately **77-92% of citations**. Compared with manual Excel-based linking, use of AI substantially reduced the time required for trial linking, achieving an **efficiency gain of approximately 89%** while maintaining reliable performance.
- Although ActiveSLR® is only one of the many tools available to streamline SLR processes, these findings suggest that the utilization of AI in the process of linking can meaningfully improve the efficiency in literature reviews, particularly in projects involving large citation volumes or complex evidence bases.
- Reviewer validation remained an essential component of the AI-supported workflow, particularly for complex publication formats such as conference abstract booklets and multi-study PDF documents. However, the time required for linking and validation was substantial lower than the manual process.
- The estimates of time saved and associated efficiency gains from AI support can have potential positive implications on evidence generation timelines. AI support can expedite the downstream workflow of a literature review which can facilitate faster regulatory and HTA submissions.

LIMITATION

- Evaluation was based on two completed SLRs. Further testing using more SLRs would be beneficial to validate these findings in more disease areas.
- Manual workflow benchmarks were derived from spreadsheet-based processes and may vary across organizations.
- ActiveSLR® is only one of many bespoke literature review programs. Comparison across different software would be beneficial for researchers, but this would require paying for access.

REFERENCES

1. Bashir R, Bourgeois FT, Dunn AG. A systematic review of the processes used to link clinical trial registrations to their published results. Syst Rev. 2017 Jul 3;6(1):123. doi: 10.1186/s13643-017-0518-3. PMID: 28669351; PMCID: PMC5494826.
2. Kwon Y, Lemieux M, McTavish J, Wathen N. Identifying and removing duplicate records from systematic review searches. J Med Libr Assoc. 2015 Oct;103(4):184-8. doi: 10.3163/1536-5050.103.4.004. PMID: 26512216; PMCID: PMC4613377.
3. Tóth B, Berek L, Gulácsi L, Péntek M, Zrubka Z. Automation of systematic reviews of biomedical literature: a scoping review of studies indexed in PubMed. Syst Rev. 2024 Jul 8;13(1):174. doi: 10.1186/s13643-024-02592-3. PMID: 38978132; PMCID: PMC11229257.
4. ActiveSLR®, Powered By PharmaQuant. Version 1.0. activeslr.io. 2025. Accessed [2026].